

fMRI-based Decoding of Visual Information from Human Brain Activity: A Brief Review

Shuo Huang^{1,2,*} Wei Shao^{1,2,*} Mei-Ling Wang^{1,2} Dao-Qiang Zhang^{1,2}

¹ College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

² MITT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing 211106, China

Abstract: One of the most significant challenges in the neuroscience community is to understand how the human brain works. Recent progress in neuroimaging techniques have validated that it is possible to decode a person's thoughts, memories, and emotions via functional magnetic resonance imaging (i.e., fMRI) since it can measure the neural activation of human brains with satisfied spatiotemporal resolutions. However, the unprecedented scale and complexity of the fMRI data have presented critical computational bottlenecks requiring new scientific analytic tools. Given the increasingly important role of machine learning in neuroscience, a great many machine learning algorithms are presented to analyze brain activities from the fMRI data. In this paper, we mainly provide a comprehensive and up-to-date review of machine learning methods for analyzing neural activities with the following three aspects, i.e., brain image functional alignment, brain activity pattern analysis, and visual stimuli reconstruction. In addition, online resources and open research problems on brain pattern analysis are also provided for the convenience of future research.

Keywords: Functional magnetic resonance imaging (fMRI), functional alignment, brain activity, brain decoding, visual stimuli reconstruction.

Citation: S. Huang, W. Shao, M. L. Wang, D. Q. Zhang. fmri-based decoding of visual information from human brain activity: a brief review. *International Journal of Automation and Computing*, vol.18, no.2, pp.170–184, 2021. <http://doi.org/10.1007/s11633-020-1263-y>

1 Introduction

One of the most significant challenges in the fields of neuroscience and machine learning is comprehending how the human brain works. As the provenance of human memory, emotion and thoughts, a better comprehension of the brain will expedite the rapid development of society, including science, medicine, education, etc.^[1–3] In order to measure neural activities, different modalities of measurement can be utilized, including event-related optical signals (EROS), positron emission tomography (PET), single-photon emission computed tomography (SPECT), near-infrared spectroscopy (NIRS), magnetoencephalography (MEG), electrocorticography (ECoG), electroencephalography (EEG), and functional magnetic resonance imaging (fMRI). Among all of the above imaging biomarkers, fMRI is one non-invasive technique for probing the neurobiological substrates of various cognitive functions that can provide indirect estimation of brain activity and measure the metabolic changes in blood

flow^[4–7]. Another advantage of fMRI is that it can provide unprecedented spatiotemporal resolution without known side effects, which intuitively can provide more accurate information for the analysis of neural activities.

Based on the fMRI images, many machine learning models are applied to analyse the visual and subjective contents of human brains^[8–10]. Generally, the machine learning-based methods aim to build a mathematical model based on the fMRI sample data, namely the training data, in order to make predictions or decisions without being explicitly programmed to perform the neural activity prediction task on the testing set. For instance, Kamitani and Tong^[11] applied a linear regression model to classify brain states and found that the cognitive trials of subjects could be reliably predicted via ensemble fMRI signals recorded in early visual areas. Kay et al.^[12] proposed a brain decoding method based on quantitative receptive field models, which learn a representation of the relationship between the stimuli images and the evoked fMRI data in early visual areas. By noticing the proportion of voxels that convey the discriminative information is small compared to the total number of measured voxels, Martino et al.^[13] applied a recursive feature elimination (RFE) algorithm to eliminate irrelevant voxels and estimate informative spatial patterns. As another work, Yamashita et al.^[14] proposed a linear classification algorithm called sparse logistic regression (SLR), that can automatically select relevant voxels as well as estimate

Review

*These authors contributed to the work equally and should be regarded as co-first authors

Manuscript received August 8, 2020; accepted October 19, 2020; published online January 16, 2021

Recommended by Associate Editor Bin Luo

Colored figures are available in the online version at <https://link.springer.com/journal/11633>

© The Author(s) 2021

their weight parameters for brain state estimation.

Although much progress has been achieved, given the data sets for the analysis of brain activities, major computational and statistical challenges have arisen to realize the full unprecedented scale and complexity of the valuable fMRI data. Overcoming these challenges has become a major and active research topic in the fields of statistical and machine learning. Here, we summarize and list the main challenges for brain pattern analysis as follows: First of all, a key component of fMRI research will be the use of multi-subject datasets. However, both anatomical structure and functional topography (brain activity patterns) vary across subjects^[15–17], and thus the authentic functional and anatomical alignments among different subjects' neural activities should be addressed before the development of the classification models. Secondly, the dimensionality of fMRI datasets is always high with redundant noise^[18, 19]. For some specific brain research experiments, such as visual or auditory stimulation, only a part of the brain area is activated in these tasks. Selecting key brain areas is a prerequisite for accurate brain research. Last but not the least, although researchers have successfully improved the classification performance for identifying brain activity patterns, the reconstruction of visual stimuli via brain images is still a challenging task^[8, 20]. Compared with the classification tasks, reconstruction of visual images can provide more detailed information for understanding human minds. In recent years, some reviews^[21–23] reviewed the mechanisms of brain encoding and decoding as well as common and

classic methods. These reviews not only summarized the up-to-date methods, but also presented the challenges in the field of brain decoding and neuroscience. In view of the above challenges, the majority of this review will be devoted to the discussion of the machine learning algorithms for solving the following four types of problems in the field of brain decoding, and we show the flowchart of our paper in Fig. 1.

Firstly, in Section 2, we will examine the problem of functional alignment for fMRI analysis across subject, which is a pre-processing step for the brain decoding analysis that takes into account variability between subjects. Since most of the research reviewed here belongs to this category, we will review a few fundamental brain alignment strategies in Section 2 including linear functional alignment, non-linear functional alignment, etc. Secondly, in Section 3, we will explore the problems of multivariate pattern classification and representation similarity analysis that predict the neural patterns with distinctive stimuli, as well as evaluate the similarities (or distances) between different cognitive tasks. Thirdly, in Section 4, we will review the methods for brain image reconstruction that generate the stimuli image via corresponding fMRI signals. Finally, online resources and open research problems on brain pattern analysis are also provided in Section 5.

2 Functional alignment

One of the challenges in the field of brain decoding is

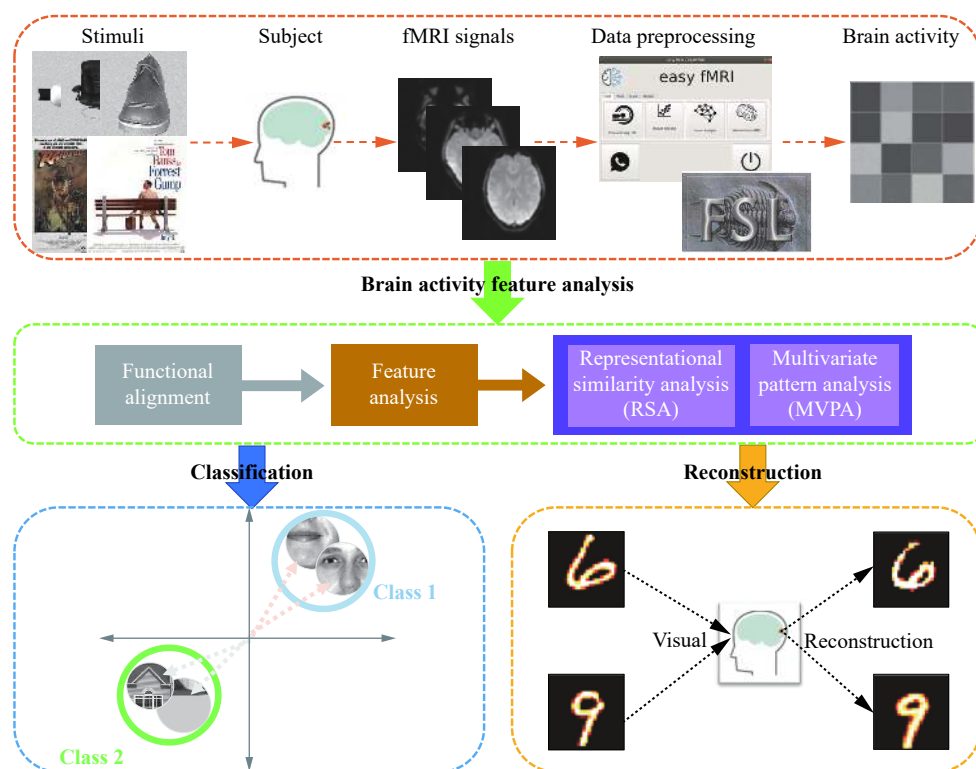


Fig. 1 General flowchart for brain decoding

the fMRI analysis of multi-subject^[1, 9, 16]. Basically, multi-subject fMRI data analysis is critical for the general evaluation of the research findings across subjects. However, due to the heterogeneous patterns in multi-subject datasets, the fMRI data collected from different subjects must be aligned into a common space in multi-subject cognitive analysis to overcome the between-subject variability^[18]. From the perspective of machine learning, we can regard the alignment problem as a multi-view representation learning problem^[1, 13]. Herein, the assumption of the alignment problem is that there is some common information across subjects, and the alignment of the data means extracting this common information. Generally, there are mainly two kinds of alignment methods, one is anatomical alignment and the other is functional alignment. The most popular method for fMRI image alignment is the anatomical alignment, which is based on anatomical features via structural MRI images, e.g., Talairach alignment^[24], or Montreal neurological institute (MNI)^[25, 26]. However, these anatomical based alignment methods cannot significantly improve the accuracy since they are insufficient to address the variability in functional topography of brains. The goal of functional alignment, on the other hand, is to precisely align the fMRI response space across the subjects. In other words, it aims at investigating a common space, where we maximize the within-class stimuli correlation and minimize the correlation between the between-class stimuli to ensure that the prominent distances exist in between-class neural activities compared with each other^[15, 16].

During the past decade, some research has combined both anatomical and functional features for fMRI functional alignment. For example, Conroy et al.^[27] proposed an alignment method that uses cortex warping to maximize the inter-subject pattern alignment. Similarly, cortical warping was used^[28] to maximize the cross-subject inter-subject correlation (ISC). As another research project focused on the maximization of ISC, Dmochowski et al.^[29] aggregated the data collected from different sub-

jects into a common matrix that can take cross-subject variability into consideration. Further, Michael et al.^[4] proposed group independent component analysis and independent vector analysis for the functional alignment of resting-state fMRI (rs-fMRI). The algorithm did not assume the simultaneity of stimuli, so it concatenated data along the temporal dimension, which means spatial consistency, and learned the components of spatial independence. Based on the above consideration, a famous alignment method, which is called hyperalignment (HA), was proposed by Haxby et al.^[1] to align the neural activity patterns across subjects onto a common space with high dimensions. Hyperalignment is a functional alignment method which is uncorrelated with anatomical features. As is shown in Fig. 2, a basic hypothesis of the original proposed hyperalignment model is that it is a common template with noisy rotations. HA uses Procrustes transform^[30] to rotate the coordinate axis of the subject's representation space, in order to align the response vectors from different subjects. The representation space of different subjects is aligned iteratively and finally a common space could be generated for all the subjects.

Followed by the work of Haxby et al.^[1], many improved alignment methods were proposed to achieve better performance. We can use different criteria to divide these methods. For example, it can be divided into supervised, semi-supervised and unsupervised functional alignment methods according to whether the label information is available. Or it can be divided into linear models and non-linear models according to the way of deriving the transform matrix. In this paper, we will introduce several classic and state-of-the-art functional alignment methods via the second division strategy.

2.1 Linear transformation methods for functional alignment

Let the matrices $X_{1:n} \in \mathbf{R}^{t \times v}$ record the data of subjects. Here, n , t and v represent the number of subjects,

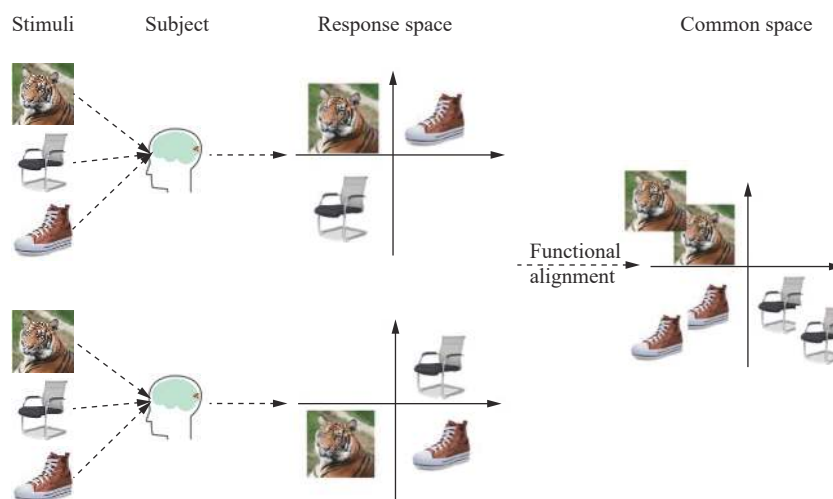


Fig. 2 Mapping different subjects into common space via Hyperalignment

the number of TRs (time of repetition) and the number of voxels, respectively. Mathematically, HA can be formulated through the framework of canonical correlation analysis (CCA)^[9]:

$$\begin{aligned} \min \quad & \sum_{i < j} \|X_i R_i - X_j R_j\|_F^2 \\ \text{s. t. } & R_k^T X_k^T X_k R_k = I, \quad k = 1, 2, \dots, n. \end{aligned} \quad (1)$$

Based on Haxby et al.'s study^[1], some studies have proposed several improved methods to ameliorate the performance of hyperalignment. Xu et al.^[31] proposed a regularized hyperalignment (RHA) method, which iteratively found the optimal regularization parameters by using the expectation-maximization (EM) algorithm. RHA proved that the weights of singular vectors in each normalized dataset are controlled by the relevant regularization parameters, and the classification accuracy can be improved by adjusting the regularization parameters. RHA verified that the weights of the singular vectors in each standardized dataset are controlled by the relevant regularization parameters, and the classification accuracy can be improved by adjusting regularization parameters. Chen et al.^[32] proposed singular value decomposition hyperalignment (SVDHA) and used joint singular value decomposition to decompose the response matrix. In this way, they reduced the dimension of fMRI for the first time. After that, HA was used to make the subjects align in a new feature space with lower dimensions, which can reduce calculation time while retaining classification accuracy. Furthermore, a shared response model (SRM) was proposed by Chen et al.^[33] as another functional alignment method. Indeed, we can think of SRM as a variant of the probability principal component analysis (PCA), and the specific way of converting is to impose orthogonal constraints on the loading matrix. One of the key attributes of SRM is the dimensional reduction mechanism, which reduces the dimensions of the shared feature space. In other studies, Sui et al.^[34, 35] applied multimodal CCA and independent component analysis (ICA) methods to multimodal data. In this way, specific and shared variance associations across multimodal data can be identified. The above studies are all based on unsupervised machine learning methods. However, in visual stimulation tasks, we can also collect supervision information such as stimulus image labels. Therefore, Yousefnezhad and Zhang^[16] proposed a supervised HA method named local discriminant hyperalignment (LDHA), which brings the concept of linear discriminant analysis (LDA) into CCA that can improve the HA performance of the unsupervised methods.

2.2 Non-linear transformation methods for functional alignment

All the HA methods mentioned above attempt to find

the transformation matrix of each subject by solving the linear model, and project the response matrices of different subjects into a common space. However, there are always nonlinearity and high-dimensionality problems in the real world. Therefore, several nonlinear HA methods were proposed for the alignment of different subjects. For example, Lorbert and Ramadge^[9] proposed a non-linear method which is called kernel hyperalignment (KHA) to do the non-linear transformation in the embedded kernel space. KHA can simultaneously solve the voxel and features expansion problems, and the difficulty of HA shifts from the limitation of the number of voxels to the number of subjects. Chen et al.^[36] developed a convolutional auto-encoder (CAE) for functional alignment on whole-brain fMRI data. As another nonlinear HA method, CAE firstly reconstructed SRM into a multi-view autoencoder. Then, CAE applied the standard searchlight (SL) to improve the stability and robustness of the cognitive classification model.

With the fast development of the deep neural networks, its powerful fitting ability provides another effective way of transformation for the nonlinear HA method. Yousefnezhad and Zhang^[17] proposed a deep hyperalignment (DHA) method as an unsupervised kernel model. As can be seen from Fig. 3, DHA used deep networks, i.e., multiple stacked layers of nonlinear transformation, as the kernel function, which can be solved via rank-m SVD and stochastic gradient descent (SGD). DHA not only solved the nonlinear problems and high-dimensional transformation, but also performed well on classification tasks.

Recently, a cross-subject graph was used by Li et al.^[15] to describe the similarities or dissimilarities among different subjects for the HA on fMRI datasets. One advantage of this method is that a new optimization algorithm based on kernels was used for nonlinear feature extraction. Here, we report the alignment results of several existing methods in Table 1, the datasets used in these methods are also shown in Table 1. More information about the presented dataset can be found in Section 5.1. It is worth noting that figures in parentheses indicate the number of categories in different datasets (ROI: region-of-interest; WB: whole-brain; PMC: post medial cortex).

3 Brain activity pattern analysis

After the functional alignment for multi-subject fMRI datasets, a common space across subjects was generated. In this new representation space, the brain activities of different subjects were represented via response matrices, in which each element denotes the activity value of a voxel or ROI. Brain activity feature analysis techniques were proposed to obtain the most discriminative features from high-dimensional, sparse, and noisy response matrices, which is also the essential prerequisite for the following classification or reconstruction task.

Research on brain activity feature analysis could be

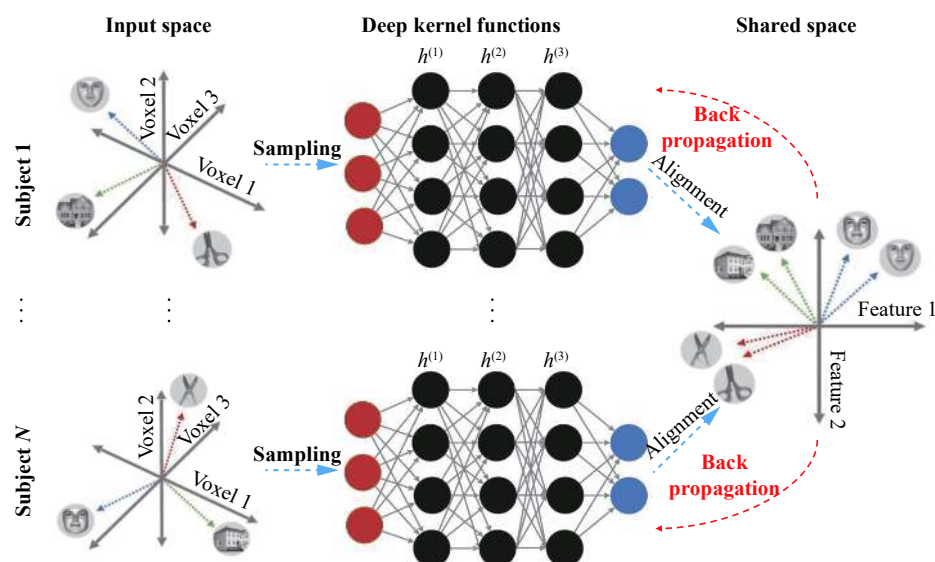
Fig. 3 Deep hyperalignment^[17]

Table 1 Performance of HA methods in post-alignment classification (%)

Methods	Datasets	Results	
		Accuracy (%)	AUC (%)
HA ^[1]	RaiderDS105	70.6±2.6 for movie time segments 63.9±2.2 for faces and objects 68.0±2.8 for animal species	
RHA ^[31]	[37]	80.85 (average performance 21% above that of basic hyperalignment)	
KHA ^[9]	Raider(7)	48.93 (Average) for ventral temporal 36.34 (Average) for entire cortex	
SRM ^[33]	Sherlock-movie	33.5±1.0	—
	Raider-movie	79.4±1.0	—
	Forrest-audio	46.5±1.0	—
	Audiobook	74.6±1.0	—
CAE ^[36]	Audiobook	18.3±1.0	—
	Sherlock-movie	42.3±1.0	—
	Movie and recall (WB)	11.8±1.0	—
	Movie and recall (PMC)	9.4±1.0	—
LDHA ^[16]	DS005(2)	94.32±0.16	93.25±0.92
	DS105ROI(8)	54.04±0.09	53.86±0.17
	DS107ROI(4)	74.73±0.19	72.03±0.37
	DS117(2)	95.07±0.27	94.23±0.94
DHA ^[17]	DS005(2)	97.92±0.82	96.91±0.82
	DS105ROI(8)	60.39±0.68	59.57±0.32
	DS107ROI(4)	73.05±0.63	70.23±0.92
	DS116(2)	90.28±0.71	89.93±0.24
Graph-based decoding model (GDM) ^[15]	DS117(2)	97.99±0.94	96.13±0.32
	DS105WB(8)	60.68±5.23	—
	DS105ROI(8)	62.22±4.23	—
	DS011(2)	92.49±2.24	—
	DS203(4)	82.47±1.45	—
	DS001(4)	62.68±1.53	—
	Raider(7)	64.52±3.28	—

traced back to 2001. Haxby et al.^[38] found that when different images are presented to a subject, different categories of visual stimuli induced different fMRI response patterns. Following their work, several brain activity pattern analysis methods^[39–42] have been proposed during the last two decades. A key concept of brain encoding and decoding is the representation of high-dimensional vector spaces. Neural responses, also known as patterns of brain activity, exist in vector form in neural representation spaces. Patterns of brain activity are distributed both spatially and temporally^[39]. Features, known as elements, in these patterns are represented as local measurements of brain activity, and each local measurement is expressed as one dimension in the space. Currently, there are numerous techniques that can be used to work with task-based fMRI datasets. These techniques, including multi-voxel pattern analysis (MVPA) and representation similarity analysis (RSA), can effectively extract and decode brain activity patterns. In this section, we will introduce the above two high-dimensional feature analysis techniques.

3.1 Multi-voxel pattern analysis

In the early days of fMRI data analysis, univariate methods were mainly used for brain activity pattern recognition. In most of these univariate methods, a general linear model (GLM)^[43] was used to estimate each voxel in the brain separately, and the analysis results were shown in an image of model parameters or derived statistics^[40]. However, with the development of research techniques, researchers found that the univariate method is not sufficient to support the analysis of fMRI data. In this case, multivariate analysis received more and more attention.

Due to the high spatial resolution of the fMRI and the particularity of the imaging method, the fMRI data have the features of high dimensionality and low signal-to-

noise ratio. The traditional univariate method treats each voxel as an independent feature, ignoring the correlation between features, which makes it difficult to detect spatial patterns^[40]. Multivariate pattern analysis, as an alternative to the traditional univariate method, can more accurately detect the activation distribution of the brain and decode the cognitive state. Therefore, multivariate pattern (MVP) analysis is widely used in many studies in the field of neuroimaging.

Information is encoded into brain activity patterns. This information comes from people's experience, or the thinking and imagination of the world. MVP analysis is a modern approach drawn from computational advances in the last two decades^[7, 41]. As one of the early studies, Haxby et al.^[1] illustrated how cognitive states can be distinguished by multi-voxel brain activity patterns. They proposed a new classifier basing on split-half correlation^[7]. The experimental results showed that a distribution representation of eight categories, such as bottles, faces, houses, etc., is contained in the ventral-temporal (VT) cortex. Furthermore, these categories could be decoded from human brain activity^[42].

Recently, sparsity learning methods have also been used to select the most discriminative voxels for brain activity pattern analysis^[5, 14, 44]. Specifically, Yamashita et al.^[14] proposed a sparse logistic regression (SLR) method, which was a linear model used for feature selection. The SLR was applied to automatically choose the most discriminative voxels in the brain and estimate the parameter weight for cognitive state identification. Moreover, Ryali et al.^[44] proposed a logistic regression-based method as well as a combination of l_1 and l_2 -norm regularizations to select discriminant brain regions across multiple conditions or groups. Grosseken et al.^[45] developed a graph-constrained elastic-net (GraphNet) based whole-brain regression and classification method that can automatically provide interpretable coefficient maps. In addition, Yousefnezhad and Zhang^[41] proposed an MVP analysis method based on the AdaBoost algorithm, which was named imbalance AdaBoost binary classification (IABC). IABC converted an imbalance MVP analysis problem to a set of balance problems to improve the fMRI analysis performance significantly. Meel et al.^[46] used MVP and functional connectivity analysis methods to study the (vertical) symmetrical representation of the regions of the ventral visual stream. Wen et al.^[5] proposed a feature selection method based on group sparse Bayesian logistic regression (GSBLR), which was applied to select the most relevant voxels for binary brain decoding. The grouped automatic relevance determination (GARD) was used in this model as prior to set the parameters, which is in concordance with the group sparsity property of the fMRI data.

3.2 Representation similarity analysis

RSA is another well-known method that is widely

used in the field of brain activity pattern analysis, which is used to evaluate the similarities between various cognitive states^[47, 48]. In a visual stimuli task, fMRI signals of subjects are acquired when watching different categories of images or videos. In a perceptual stimuli task, different categories of stimuli can evoke corresponding activity patterns in the brain of a subject. Then, RSA will be used to calculate the similarities between various cognitive states. This process will generate the representational similarity matrix (RSM) that encodes the similarity structure of different cognitive tasks. Fig. 4 shows the computational steps for the derivation of the RSM. In the RSM, each block represents a correlation distance between the activity patterns of a pair of stimuli (i.e., conditions in the experiment). The diagonal elements of the RSM are equal to 1. The value of matrix's non-diagonal elements represents the similarity of brain's responses to two different stimuli. The larger the value, the higher the similarity, vice versa.

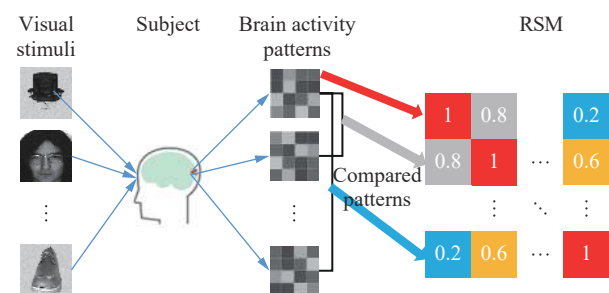


Fig. 4 Computation of RSM

Classic RSA is mainly based on traditional linear methods, e.g., GLM^[43], ordinary least squares (OLS)^[47], etc. In fact, we can regard RSA as a multi-task regression problem. Kriegeskorte et al.^[47] used the ordinary least squares method to fit the linear model of the time frame for each voxel to measure the spatial activity patterns caused in each condition. This linear model includes a hemodynamic response predictor for each case, as well as an optional further predictor for modeling human factors, such as trends, head movement effects, and baseline shifts between measurement runs. RSA^[48] assumes that the brain activity patterns are related to stimuli events, which can be formulated as

$$\mathbf{Y}^{(i)} = \mathbf{X}^{(i)} \mathbf{B}^{(i)} + \boldsymbol{\epsilon}^{(i)} \quad (2)$$

where $\mathbf{Y}^{(i)} = \{y_{mn}\} \in \mathbf{R}^{T \times V}$, $1 \leq m \leq T$, $1 \leq n \leq V$, denotes the fMRI time series from the i -th subject, T is the number of repetition time (TR) and V is the number of voxels of brain. The design matrix is denoted by $\mathbf{X}^{(i)} = \{x_{mk}\} \in \mathbf{R}^{T \times P}$, $1 \leq m \leq T$, $1 \leq k \leq P$. The design matrix $\mathbf{X}^{(i)}$ can be obtained by the convolution of the stimuli time series with a typical hemodynamic response function (HRF). Here, P denotes the number of the categories of stimuli, $\mathbf{B}^{(i)} = \{\beta_{kj}\} \in \mathbf{R}^{P \times V}$, $\beta_{kj} \in \mathbf{R}$, $1 \leq k \leq P$, $1 \leq j \leq V$,

denotes the estimated regression matrix, and β_{kj} is an amplitude reflecting the response of the j -th voxel to the k -th stimuli. GLM is based on a linear model and it cannot achieve satisfactory results since the representation matrix is usually a wide matrix, which means that the voxel account is far more than the time points in fMRI dataset. Moreover, this method makes it difficult to convert data into a matrix^[41]. Also, the method's stability and robustness will decrease when the value of signal-to-noise (SNR) reduces^[2]. Further, GLM and OLS will face the problem of overfitting. Most of the existing studies avoid overfitting by adding the regularization terms. For instance, the least absolute shrinkage and selection operator (LASSO)^[49] was proposed to solve the regression problem by using l_1 -norm, whereas l_2 -norm was used in the ridge regression^[50] method to address the aforementioned problem. The elastic net^[51], as a modified model, was developed to address the above issues via combining l_1 and l_2 norms.

On the other hand, a concept called searchlight was introduced by researchers as an alternative method of region-of-interest (ROI) based fMRI analysis. SL implements MVP analysis on sphere-shaped groups of voxels centered on each voxel one by one^[5]. As we mentioned before, due to the high spatial resolution of fMRI data, the whole-brain datasets have high dimensionality. In the past, when using RSA methods, it was difficult to convert the data into a matrix and we could not avoid the inverse of the voxel matrix. In addition, when the number of voxels is too large, RSA optimization is also plagued by high-dimension data. Fortunately, compared with traditional RSA algorithms, modern RSA algorithms can optimize the solution process^[52]. Su et al.^[53] proposed an RSA method that uses searchlight technology for EMEG (a combination of MEG and EEG). This method directly implemented the MVP analysis of information flow in the human brain and the spatial and temporal identification of fine-grained dynamic neural

calculations. As an extended application, the SL-based RSA method can also be applied for the structure analysis in the ethical violation space^[54].

In short, RSA provides researchers with a new perspective to compare different genomic representation across different subjects, different ROI from one subject, different modalities of measurement, and even different species. Since similarity structures can be estimated from imaging data even without coding models, RSA cannot only be used for model testing but also for exploratory research^[48]. RSA is also initially used to study visual representations^[43, 55, 56], semantic representations^[52, 57] and lexical representations^[53]. Last but not the least, RSA can also be applied to reveal the representations of social networks^[58, 59].

4 Visual stimuli reconstruction

Like the classification and regression task in machine learning, the purpose of brain decoding is to analyze the subject's brain activity patterns to perform the task of visual stimuli identification or reconstructing the stimuli details. In recent years, quite a lot of studies have been made for the classification of brain activity patterns^[1, 38, 45]. However, the reconstruction of brain images is still a challenging task. A general conceptual framework for visual stimuli reconstruction is shown in Fig. 5, which can be regarded as a cross-modal reconstruction (The green line represents image reconstruction while the blue line denotes the fMRI). Visual stimuli reconstruction focuses on acquiring the relevant features between the stimuli images and fMRI in order to generate the stimuli images via the corresponding fMRI signal.

Many researchers have made preliminary explorations in the field of visual stimuli reconstruction. As an early exploratory study, Thirion et al.^[60] used rotating Gabors to reconstruct dot patterns from stimuli and imagery. They predicted the visual stimuli of both real and imagin-

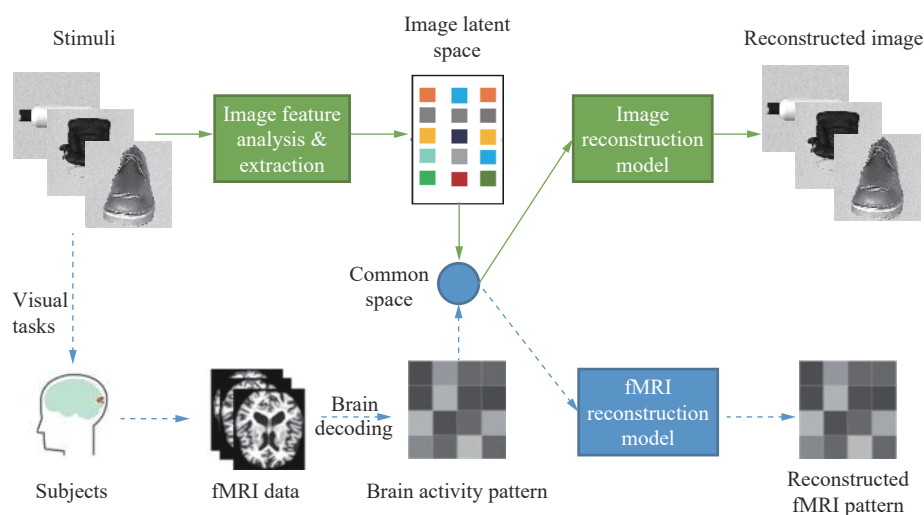


Fig. 5 General conceptual framework for visual stimulation reconstruction

ary scenes via the evoked brain activities, which was elicited from the visual cortex. Moreover, Miyawaki et al.^[61] firstly asked the volunteers to watch a lot of flashing checker board images as visual stimuli and recorded the evoked brain activity patterns of these stimuli in the early visual cortex (V1/V2/V3) and then built a sparse multi-scale multinomial logistic regression (SMLR) local decoder model for visual stimuli reconstruction. The experimental results showed that this method provided a new way to interpret the visual perception of the brain.

In recent years, many reconstruction methods have been proposed for visual stimuli reconstruction. These methods can be divided into traditional machine learning methods and the latest deep network framework. Among the traditional machine learning methods, the Bayesian model is the most common one. In this paper, we will review the recent progress with the following two aspects, i.e., the Bayesian-based reconstruction models and deep generation model-based reconstruction methods.

4.1 Bayesian-based reconstruction model

Inspired by the work of Miyawaki et al.^[61], some reconstruction models based on Bayesian models are proposed to explore the correlations among the signals recorded in fMRI that can reflect the features of corresponding stimuli images. For example, Naselaris et al.^[62] proposed a joint model that combines structural and semantic features of brain activity patterns. And a Bayesian framework is used here to infer the stimuli images from a large-scale dataset via the evoked brain activities. Nishimoto et al.^[63] used a Bayesian decoding framework for movie scene reconstruction from the given blood-oxygen-level-dependent (BOLD) signals. A motion-energy encoding model is proposed by the authors that largely overcomes the limitation of tardiness of BOLD signals measured via fMRI. Further, a model called Bayesian canonical correlation analysis (BCCA) was proposed by Fujiwara et al.^[64] to automatically learn image bases. CCA was used to construct an invertible mapping based on the Bayesian model. Zhan et al.^[10] proposed a reconstruction method based on a support vector machine (SVM) and Bayesian classifier followed by ICA to improve the efficiency of feature extraction and reconstruction performance. Cowen et al.^[65] used PCA to transform human face stimuli into a new feature space, and then established the relationship between new features and fMRI signals, and realized reconstruction of human face stimuli for the first time. Du et al.^[66] proposed a Bayesian-based reconstruction method that derives missing latent variables by Bayesian inference. The joint generative model of external stimuli and brain activities they proposed can not only extract non-linear features of the stimuli images, but also capture the correlation among brain activities. The reconstruction models based on the Bayesian framework aims to find the relationship between the visual

stimuli and the corresponding fMRI signals, and establish a linear mapping between them to achieve the task of image reconstruction. However, the linear mapping often cannot truly reflect the relationship between the two cross-modal data, and the reconstruction results obtained are often coarse-grained, making it difficult to describe the details of the images.

4.2 Deep network-based reconstruction model

In the last decade, deep learning has drawn significant attentions for its powerful fitting and generating capabilities. Variational autoencoder (VAE)^[67] and generative adversarial network (GAN)^[68] are two of the most popular approaches. VAE describes potential spatial observations in a probabilistic manner. Therefore, instead of constructing an encoder that outputs a single value to describe each latent state attribute, we use an encoder to describe the probability distribution of each latent attribute. By sampling from the underlying space, we can use the network of decoders to form a generative model that can create new data that are like the observations of the training data. In other words, we could sample from the prior distribution $p(z)$, and assume that it follows a unit Gaussian distribution. Recently, Du et al.^[8] proposed a deep generative multi-view model (DGMM) for stimuli image reconstruction from the evoked brain activity patterns. DGMM can be regarded as a nonlinear extension of BCCA by combining image generation models with Bayesian inferences to accomplish reconstruction tasks.

As other deep learning approaches, Horikawa and Kamitani^[69] presented a brain decoding method via the computer vision principle, which represent the categories with a group of latent features through hierarchical processing. By this way, they found that the features of visual images can be predicted from brain activities of subjects. A model based on a deep neural network (DNN) was trained by Shen et al.^[20] to establish an end-to-end reconstruction model via visual stimuli images and the evoked brain activity patterns. Experimental results showed that a direct mapping can be learned by the proposed model for perceptual reconstruction.

GAN is another relatively important model in the field of deep learning. The original GAN model is proposed by Goodfellow et al.^[68] in 2014, where the discriminator and generator play the following mini-max game:

$$\min_{\beta} \max_{\theta} F(\beta, \theta) = E_{x \sim p_{data}} [\log_{p_{\theta}} (y = 1|x)] + E_{x \sim p_{\beta}} [\log_{p_{\theta}} (y = 0|x)] \quad (3)$$

where p_{data} is the distribution of real data, p_{σ} is the distribution of generated data, and p_{θ} is the distribution of the discriminator with parameter θ , E represents the mathematical expression. The training stage of GAN can

be seen as a zero-sum game, in which the generator tries to generate the data that can fool the discriminator, and the discriminator is used to distinguish the generated fake data y from the real data x and label them with 1 and 0, respectively.

Some GAN-based visual stimuli reconstruction models have been proposed and greatly improved the precision of the reconstruction results. For instance, St-Yves and Naselaris^[70] used GAN architecture to learn an image generation model and completed perceptual stimuli reconstruction through this model. And in this way, the noise model can be inferred from the measured brain activity. Furthermore, some approaches based on GANs are proposed to reconstruct human face images. Güçlütürk et al.^[71] proposed a joint model to combine probabilistic inference with the GAN architecture for face stimuli reconstruction from human brain activities. They maximized posteriori estimation to invert the linear transformation from features in latent space to brain activity patterns. Then, the convolutional neural networks (CNN) were used to invert a non-linear transformation from visual stimuli to latent features. Seeliger et al.^[72] introduced a deep convolutional generative adversarial network (DCGAN) architecture to reconstruct the stimuli images. Also, they used a linear model to predict the latent space of a generative model from the evoked brain activity patterns. More recently, VanRullen and Reddy^[73] presented thousands of celebrity face images of a large dataset to the subjects as a stimuli task. Then, they trained a VAE neural network using a GAN architecture over this dataset and learnt a linear mapping between

face images and fMRI activity patterns. Compared with the classic linear reconstruction methods, models based on deep networks can implement non-linear transformations that greatly improve the accuracy of image reconstruction, and describe images in fine granularity. Fig. 6 shows some experimental results of several visual stimuli reconstruction tasks in recent years. In addition to the visual stimuli reconstruction methods based on Bayesian or deep neural networks we mentioned above, considering it is difficult to collect a large amount of pairwise image-fMRI data for training, there are several methods^[74–78] using semi-supervised learning (SSL) to improve brain decoding performance by leveraging large number of images.

5 Open resources & future work

5.1 Open resources

As we all know, the collection of high-quality datasets is an important guarantee for the research of data-driven machine learning methods. For the decoding of visual information from human brain activity, Open NEURO¹ project is a free and open platform for sharing MRI, MEG, EEG, iEEG and ECoG data. As an extended version of Open fMRI project, the project now has 404 available datasets and 12 037 participants across all datasets. Table 2 shows some datasets in Open NEURO project.

In order to promote the rapid development of the field

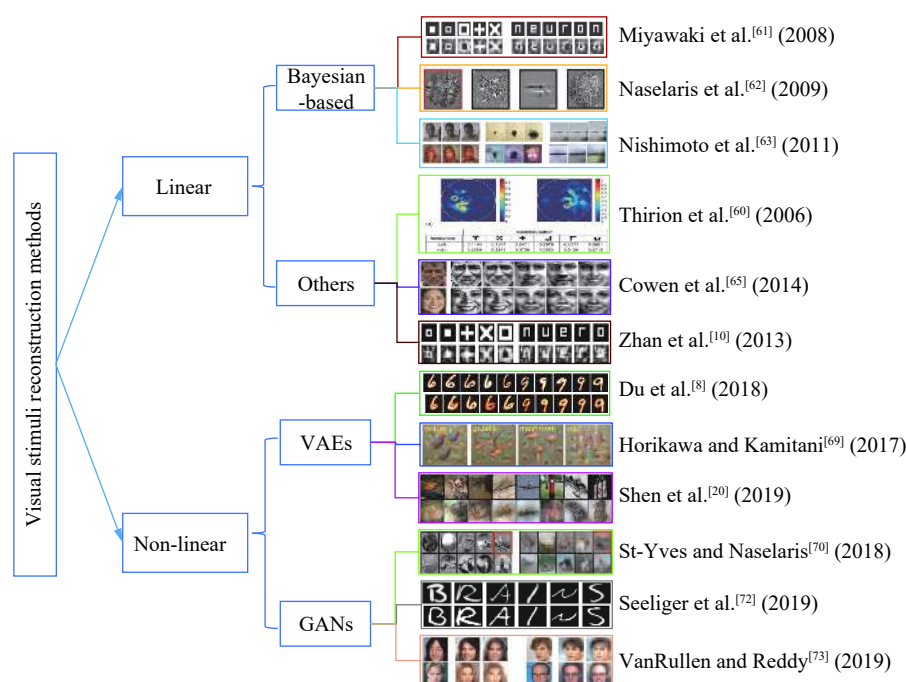


Fig. 6 A brief presentation of the results of some visual stimuli reconstruction methods in recent years

¹<http://openneuro.org/>

Table 2 Dataset descriptions in Open NEURO Project

ID	Titles	Subjects	Categories	Time points	TR	TE	Ref.
DS001	Balloon Analog Risk	16	4	894	2000	77	[79]
DS002D	Deterministic classification	17	2	356	2000	20	[80]
DS002P	Probabilistic classification	17	2	356	2000	20	[80]
DS005	Mixed-gambles	16	4	714	2000	30	[81]
DS011D	Dual-task weather prediction	14	3	408	2000	25	[82]
DS011W	Weather prediction without feedback	14	4	236	2000	25	[82]
DS017	Selective stop signal task	8	6	546	2000	25	[80]
DS052R	Reversal weather prediction	13	2	450	2000	20	[83]
DS052W	Weather prediction	13	2	450	2000	20	[83]
DS102	Flanker task	26	2	292	2000	20	[84]
DS105	Visual object recognition	6	8	1 452	2 500	30	[38]
DS107	Word and object processing	49	4	322	2000	28	[85]
DS116A	Auditory odd ball	17	2	510	2000	20	[86]
DS116V	Visual odd ball	17	2	510	2000	20	[86]
DS164	Stroop	28	2	370	1 500	10	[87]
DS231	Integration of sweet taste	9	6	1 119	2000	30	[88]
DS232	Face-coding localizer (objects) task	10	4	760	1 060	16	[47]

TR: Repetition time in millisecond; TE: Echo time in millisecond.

of brain science and neural computing, more and more researchers have made their works open source online. These works progress mainly includes algorithms and the open-source Software packages. For example, Chen et al. [19, 32, 33] made their codes on brain pattern analysis available online, including some proposed algorithms and open-source libraries such as SciKit-Learn for model training. What's more, some research groups developed open source software are for brain image analysis. One of the most famous examples is PyMVPA² developed by the Haxby Lab³ at Dartmouth College. PyMVPA is an open source software toolbox based on Python, which is used for the application of analysis techniques based on classifiers to fMRI datasets. PyMVPA is a cross-platform toolbox that makes use of the abilities of Python to access the libraries which are written in various of programming languages and computing environments to interface with the wealth of existing machine learning packages [89, 90].

Recently, a new toolbox called easy fMRI⁴ is developed for analyzing fMRI datasets (shown in Fig. 7). Easy fMRI is a toolbox with the capability of decoding and visualizing the human brain. It is developed by the iBRAIN⁵ research group of Nanjing University of Aeronautics and Astronautics, which is free and open source. It is designed based on the brain imaging data structure (BIDS) file, which supports automatic labelling on the de-

signed matrix.

Easy fMRI uses advanced machine learning techniques and high-performance computing to analyze task-based fMRI datasets. It provides a friendly graphical user interface for feature analysis, HA, MVPA, RSA, etc. In addition, easy fMRI is integrated with FMRIB Software Library (for the preprocessing step), SciKit-Learn (for model analysis), PyTorch (for deep learning methods), and AFNI/SUMA (for 3D visualization). ANFI represents analysis of functional neuroimages. SUMA allows viewing 3D cortical surface model and mapping volumetric data onto them.

5.2 Future work

In this paper, we reviewed the methods developed and employed for the decoding of visual information from human brain activity. In future studies, there are several issues needing to be addressed. For instance, task-based fMRI is difficult to collect due to the difficulty in keeping subjects' heads stationary. Therefore, the sample sizes of most task-based fMRI datasets are small. Some studies [91–93] are proposed to process them by applying domain adaptation and transfer learning algorithms. In order to make the brain decoding algorithms available to large scale and multi-site fMRI datasets, this is an important issue and needs more studies.

Furthermore, the three aspects we mentioned above, i.e., brain image alignment, brain decoding and brain image reconstruction are usually studied independently. In the future, we will consider combining them together to deal with more complex real-world problems. For ex-

²<http://www.pympva.org/>

³<http://haxbylab.dartmouth.edu/>

⁴<https://easyfmri.learningbymachine.com/>

⁵<http://ibrain.nuaa.edu.cn/>

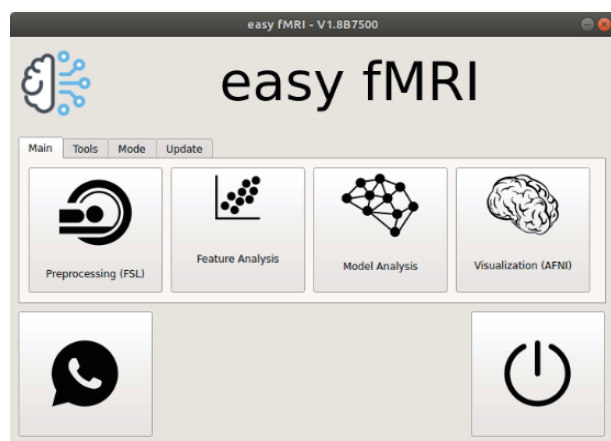


Fig. 7 A screenshot of easy fMRI toolbox

ample, Du et al.^[8] mentioned that in the visual stimuli reconstruction task, the reconstruction results of different subjects were significantly different. To solve this problem, we can combine HA and the reconstruction task to reduce reconstruction differences across subjects.

Finally, most of the current methods do not make good use of the structural information of the whole brain structure data. In future studies, we plan to develop information-based models on the basis of understanding the intrinsic information of the whole brain structure data to smooth the data information of small areas. It makes the information valid area in the whole brain data clearer and provides better input information for subsequent feature selection and representation similarity analysis.

6 Conclusions

In this paper, we have reviewed the mechanisms and the strategies of machine learning methods for analyzing neural activities via fMRI data. As an interdisciplinary field of research, computational neuroscience can break the neural codes via different concepts from different subjects such as mathematics, psychology, machine learning, etc. However, there are still some challenges in the field of fMRI research such as multi-subject datasets, high-dimensional feature analysis and the generation of visual images from fMRI. We conducted a brief review on the state-of-the-art machine learning techniques for solving these challenges, including linear and nonlinear functional alignment, multi-voxel pattern analysis, representation similarity analysis and visual stimuli reconstruction based on Bayesian or deep neural networks. Last but not least, we also provided online resources and open research problems on brain pattern analysis for the convenience of future research, and put forward some ideas for future work in the field of brain science and neural computing.

Acknowledgements

This work was supported by National Natural Science

Foundation of China (Nos. 61876082, 61861130366, 6173-2006 and 61902183), National Key Research and Development Program of China (Nos. 2018 YFC2001600, 2018YFC 2001602), the Royal Society-Academy of Medical Sciences Newton Advanced Fellowship (No. NAF\R1\180371), and China Postdoctoral Science Foundation funded project (No. 2019M661831).

Open Access

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- [1] J. V. Haxby, J. S. Guntupalli, A. C. Connolly, Y. O. Halchenko, B. R. Conroy, M. I. Gobbini, M. Hanke, P. J. Ramadge. A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron*, vol. 72, no. 2, pp. 404–416, 2011. DOI: [10.1016/j.neuron.2011.08.026](https://doi.org/10.1016/j.neuron.2011.08.026).
- [2] M. B. Cai, N. W. Schuck, J. W. Pillow, Y. Niv. A Bayesian method for reducing bias in neural representational similarity analysis. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Barcelona, Spain, pp. 4958–4966, 2016.
- [3] J. H. Tao, J. Huang, Y. Li, Z. Lian, M. Y. Niu. Semi-supervised ladder networks for speech emotion recognition. *International Journal of Automation and Computing*, vol. 16, no. 4, pp. 437–448, 2019. DOI: [10.1007/s11633-019-1175-x](https://doi.org/10.1007/s11633-019-1175-x).
- [4] A. M. Michael, M. Anderson, R. L. Miller, T. Adalı, V. D. Calhoun. Preserving subject variability in group fMRI analysis: Performance evaluation of GICA vs. IVA. *Frontiers in Systems Neuroscience*, vol. 8, Article number 106, 2014. DOI: [10.3389/fnsys.2014.00106](https://doi.org/10.3389/fnsys.2014.00106).
- [5] Z. F. Wen, T. Y. Yu, Z. L. Yu, Y. Q. Li. Grouped sparse Bayesian learning for voxel selection in multivoxel pattern analysis of fMRI data. *NeuroImage*, vol. 184, pp. 417–430, 2019. DOI: [10.1016/j.neuroimage.2018.09.031](https://doi.org/10.1016/j.neuroimage.2018.09.031).
- [6] D. Haputhanthri, G. Brihadiswaran, S. Gunathilaka, D. Meedeniya, S. Jayarathna, M. Jaime, C. Harshaw. Integration of facial thermography in EEG-based classification of ASD. *International Journal of Automation and Computing*, to be published. DOI: [10.1007/s11633-020-1231-6](https://doi.org/10.1007/s11633-020-1231-6).
- [7] J. V. Haxby. Multivariate pattern analysis of fMRI: The

- early beginnings. *NeuroImage*, vol. 62, no. 2, pp. 852–855, 2012. DOI: [10.1016/j.neuroimage.2012.03.016](https://doi.org/10.1016/j.neuroimage.2012.03.016).
- [8] C. D. Du, C. Y. Du, L. J. Huang, H. G. He. Reconstructing perceived images from human brain activities with Bayesian deep multiview learning. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 8, pp. 2310–2323, 2019. DOI: [10.1109/TNNLS.2018.2882456](https://doi.org/10.1109/TNNLS.2018.2882456).
- [9] A. Lorbert, P. J. Ramadge. Kernel hyperalignment. In *Proceedings of the 25th International Conference on Neural Information Processing Systems* Lake Tahoe, USA, pp. 1790–1798, 2012.
- [10] Y. Zhan, J. C. Zhang, S. T. Song, L. Yao. Visual image reconstruction from fMRI activation using multi-scale support vector machine decoders. In *Proceedings of the 15th International Conference on Human-Computer Interaction*, Springer, Las Vegas, USA, pp. 491–497, 2013. DOI: [10.1007/978-3-642-39342-6_54](https://doi.org/10.1007/978-3-642-39342-6_54).
- [11] Y. Kamitani, F. Tong. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, vol. 8, no. 5, pp. 679–685, 2005. DOI: [10.1038/nn1444](https://doi.org/10.1038/nn1444).
- [12] K. N. Kay, T. Naselaris, R. J. Prenger, J. L. Gallant. Identifying natural images from human brain activity. *Nature*, vol. 452, no. 7185, pp. 352–355, 2008. DOI: [10.1038/nature06713](https://doi.org/10.1038/nature06713).
- [13] F. De Martino, G. Valente, N. Staeren, J. Ashburner, R. Goebel, E. Formisano. Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. *NeuroImage*, vol. 43, no. 1, pp. 44–58, 2008. DOI: [10.1016/j.neuroimage.2008.06.037](https://doi.org/10.1016/j.neuroimage.2008.06.037).
- [14] O. Yamashita, M. A. Sato, T. Yoshioka, F. Tong, Y. Kamitani. Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *NeuroImage*, vol. 42, no. 4, pp. 1414–1429, 2008. DOI: [10.1016/j.neuroimage.2008.05.050](https://doi.org/10.1016/j.neuroimage.2008.05.050).
- [15] W. D. Li, M. X. Liu, F. Chen, D. Q. Zhang. Graph-based decoding model for functional alignment of unaligned fMRI data. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, New York, USA, pp. 2653–2660, 2020.
- [16] M. Yousefnezhad, D. Q. Zhang. Local discriminant hyperalignment for multi-subject fMRI data alignment. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, California, USA, pp. 59–65, 2017. DOI: [10.1016/j.neuroimage.2008.05.050](https://doi.org/10.1016/j.neuroimage.2008.05.050).
- [17] M. Yousefnezhad, D. Q. Zhang. Deep hyperalignment. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, USA, pp. 1603–1611, 2017.
- [18] J. S. Guntupalli, M. Hanke, Y. O. Halchenko, A. C. Connolly, P. J. Ramadge, J. V. Haxby. A model of representational spaces in human cortex. *Cerebral Cortex*, vol. 26, pp. 2919–2934, 2016. DOI: [10.1093/cercor/bhw068](https://doi.org/10.1093/cercor/bhw068).
- [19] P. H. Chen. Multi-view Representation Learning with Applications to Functional Neuroimaging Data, Princeton University, Ph. D. dissertation, USA, 2017.
- [20] G. H. Shen, K. Dwivedi, K. Majima, T. Horikawa, Y. Kamitani. End-to-end deep image reconstruction from human brain activity. *Frontiers in Computational Neuroscience*, vol. 13, Article number 21, 2019. DOI: [10.3389/fncom.2019.00021](https://doi.org/10.3389/fncom.2019.00021).
- [21] T. Naselaris, K. N. Kay, S. Nishimoto, J. L. Gallant. Encoding and decoding in fMRI. *NeuroImage*, vol. 56, no. 2, pp. 400–410, 2011. DOI: [10.1016/j.neuroimage.2010.07.073](https://doi.org/10.1016/j.neuroimage.2010.07.073).
- [22] C. D. Du, J. P. Li, L. J. Huang, H. G. He. Brain encoding and decoding in fMRI with bidirectional deep generative models. *Engineering*, vol. 5, no. 5, pp. 948–953, 2019. DOI: [10.1016/j.eng.2019.03.010](https://doi.org/10.1016/j.eng.2019.03.010).
- [23] K. Vakamudi, S. Posse, R. Jung, B. Cushman, M. O. Chohan. Real-time presurgical resting-state fMRI in patients with brain tumors: Quality control and comparison with task-fMRI and intraoperative mapping. *Human Brain Mapping*, vol. 41, pp. 797–814, 2020. DOI: [10.1002/hbm.24840](https://doi.org/10.1002/hbm.24840).
- [24] J. Talairach P. Tournoux. 3-dimensional proportional system: An approach to cerebral imaging. Co-Planar Stereotaxic Atlas of the Human Brain. *Thieme*, 1988.
- [25] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proceedings of IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, IEEE, San Francisco, USA, pp. 1813–1817, 1993. DOI: [10.1109/NSSMIC.1993.373602](https://doi.org/10.1109/NSSMIC.1993.373602).
- [26] W. Chau, A. R. McIntosh. The Talairach coordinate of a point in the MNI space: How to interpret it. *NeuroImage*, vol. 25, no. 2, pp. 408–416, 2005. DOI: [10.1016/j.neuroimage.2004.12.007](https://doi.org/10.1016/j.neuroimage.2004.12.007).
- [27] B. R. Conroy, B. D. Singer, J. V. Haxby, P. J. Ramadge. fMRI-based inter-subject cortical alignment using functional connectivity. In *Proceedings of the 22nd International Conference on Neural Information Processing Systems*, Vancouver, Canada, pp. 378–386, 2009.
- [28] M. R. Sabuncu, B. D. Singer, B. Conroy, R. E. Bryan, P. J. Ramadge, J. V. Haxby. Function-based intersubject alignment of human cortical anatomy. *Cerebral Cortex*, vol. 20, no. 1, pp. 130–140, 2009. DOI: [10.1093/cercor/bhp085](https://doi.org/10.1093/cercor/bhp085).
- [29] J. P. Dmochowski, P. Sajda, J. Dias, L. C. Parra. Correlated components of ongoing EEG point to emotionally laden attention—a possible marker of engagement? *Frontiers in Human Neuroscience*, vol. 6, Article number 112, 2012. DOI: [10.3389/fnhum.2012.00112](https://doi.org/10.3389/fnhum.2012.00112).
- [30] P. H. Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, vol. 31, no. 1, pp. 1–10, 1966. DOI: [10.1007/BF02289451](https://doi.org/10.1007/BF02289451).
- [31] H. Xu, A. Lorbert, P. J. Ramadge, J. S. Guntupalli, J. V. Haxby. Regularized hyperalignment of multi-set fMRI data. In *Proceedings of IEEE Statistical Signal Processing Workshop*, IEEE, Ann Arbor, USA, pp. 229–232, 2012. DOI: [10.1109/SSP.2012.6319668](https://doi.org/10.1109/SSP.2012.6319668).
- [32] P. H. Chen, J. S. Guntupalli, J. V. Haxby, P. J. Ramadge. Joint SVD-Hyperalignment for multi-subject fMRI data alignment. In *Proceedings of IEEE International Workshop on Machine Learning for Signal Processing*, IEEE, Reims, France, 2014. DOI: [10.1109/MLSP.2014.6958912](https://doi.org/10.1109/MLSP.2014.6958912).
- [33] P. H. Chen, J. Chen, Y. Yeshurun, U. Hasson, J. V. Haxby, P. J. Ramadge. A reduced-dimension fMRI shared response model. In *Proceedings of the 28th International Conference on Neural Information Processing Systems*, Montreal, Canada, pp. 460–468, 2015.
- [34] J. Sui, G. Pearlson, A. Caprihan, T. Adali, K. A. Kiehl, J. Y. Liu, J. Yamamoto, V. D. Calhoun. Discriminating schizophrenia and bipolar disorder by fusing fMRI and DTI in a multimodal CCA+ joint ICA model. *NeuroImage*, vol. 57, no. 3, pp. 839–855, 2011. DOI: [10.1016/j.neuroimage.2011.05.055](https://doi.org/10.1016/j.neuroimage.2011.05.055).
- [35] J. Sui, H. He, G. D. Pearlson, T. Adali, K. A. Kiehl, Q. B. Yu, V. P. Clark, E. Castro, T. White, B. A. Mueller, B. C. Ho, N. C. Andreasen, V. D. Calhoun. Three-way (N-way)

- fusion of brain imaging data based on mCCA+ jICA and its application to discriminating schizophrenia. *NeuroImage*, vol. 66, pp. 119–132, 2013. DOI: [10.1016/j.neuroimage.2012.10.051](https://doi.org/10.1016/j.neuroimage.2012.10.051).
- [36] P. H. Chen, X. Zhu, H. J. Zhang, J. S. Turek, J. Chen, T. L. Willke, U. Hasson, P. J. Ramadge. A convolutional autoencoder for multi-subject fMRI data aggregation. [Online], Available: <https://arxiv.org/abs/1608.04846>, 2016.
- [37] U. Hasson, O. Landesman, B. Knappmeyer, I. Vallines, N. Rubin, D. J. Heeger. Neurocinematics: The neuroscience of film. *Projections*, vol. 2, no. 1, pp. 1–26, 2008. DOI: [10.3167/proj.2008.020102](https://doi.org/10.3167/proj.2008.020102).
- [38] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, P. Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, vol. 293, no. 5539, pp. 2425–2430, 2001. DOI: [10.1126/science.1063736](https://doi.org/10.1126/science.1063736).
- [39] J. V. Haxby, A. C. Connolly, J. S. Guntupalli. Decoding neural representational spaces using multivariate pattern analysis. *Annual Review of Neuroscience*, vol. 37, pp. 435–456, 2014. DOI: [10.1146/annurev-neuro-062012-170325](https://doi.org/10.1146/annurev-neuro-062012-170325).
- [40] C. Allefeld, J. D. Haynes. Multi-voxel pattern analysis. *Brain Mapping*, vol. 1, pp. 641–646, 2015. DOI: [10.1016/B978-0-12-397025-1.00345-6](https://doi.org/10.1016/B978-0-12-397025-1.00345-6).
- [41] M. Yousefnezhad, D. Q. Zhang. Anatomical pattern analysis for decoding visual stimuli in human brains. *Cognitive Computation*, vol. 10, no. 2, pp. 284–295, 2018. DOI: [10.1007/s12559-017-9518-9](https://doi.org/10.1007/s12559-017-9518-9).
- [42] D. D. Wagner, R. S. Chavez, T. W. Broom. Decoding the neural representation of self and person knowledge with multivariate pattern analysis and data-driven approaches. *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 10, no. 1, Article number e1482, 2019. DOI: [10.1002/wcs.1482](https://doi.org/10.1002/wcs.1482).
- [43] A. C. Connolly, J. S. Guntupalli, J. Gors, M. Hanke, Y. O. Halchenko, Y. C. Wu, H. Abdi, J. V. Haxby. The representation of biological classes in the human brain. *Journal of Neuroscience*, vol. 32, no. 8, pp. 2608–2618, 2012. DOI: [10.1523/JNEUROSCI.5547-11.2012](https://doi.org/10.1523/JNEUROSCI.5547-11.2012).
- [44] S. Ryali, K. Supekar, D. A. Abrams, V. Menon. Sparse logistic regression for whole-brain classification of fMRI data. *NeuroImage*, vol. 51, no. 2, pp. 752–764, 2010. DOI: [10.1016/j.neuroimage.2010.02.040](https://doi.org/10.1016/j.neuroimage.2010.02.040).
- [45] L. Grosse, B. Klingenberg, S. Greer, J. Taylor, B. Knutson. Whole-brain sparse penalized discriminant analysis for predicting choice. *NeuroImage*, vol. 47, no. S1, Article number S58, 2009. DOI: [10.1016/S1053-8119\(09\)70232-0](https://doi.org/10.1016/S1053-8119(09)70232-0).
- [46] C. van Meel, A. Baeck, C. R. Gillebert, J. Wagemans, H. P. O. de Beeck. The representation of symmetry in multi-voxel response patterns and functional connectivity throughout the ventral visual stream. *NeuroImage*, vol. 191, pp. 216–224, 2019. DOI: [10.1016/j.neuroimage.2019.02.030](https://doi.org/10.1016/j.neuroimage.2019.02.030).
- [47] N. Kriegeskorte, R. Goebel, P. Bandettini. Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 10, pp. 3863–3868, 2006. DOI: [10.1073/pnas.0600244103](https://doi.org/10.1073/pnas.0600244103).
- [48] N. Kriegeskorte, M. Mur, P. Bandettini. Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, vol. 2, Article number 4, 2008. DOI: [10.3389/neuro.06.004.2008](https://doi.org/10.3389/neuro.06.004.2008).
- [49] M. Yuan, Y. Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006. DOI: [10.1111/j.1467-9868.2005.00532.x](https://doi.org/10.1111/j.1467-9868.2005.00532.x).
- [50] A. E. Hoerl, R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, vol. 42, no. 1, pp. 80–86, 2000. DOI: [10.1080/00401706.2000.10485983](https://doi.org/10.1080/00401706.2000.10485983).
- [51] H. Zou, T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, 2005. DOI: [10.1111/j.1467-9868.2005.00503.x](https://doi.org/10.1111/j.1467-9868.2005.00503.x).
- [52] A. G. Huth, S. Nishimoto, A. T. Vu, J. L. Gallant. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, vol. 76, no. 6, pp. 1210–1224, 2012. DOI: [10.1016/j.neuron.2012.10.014](https://doi.org/10.1016/j.neuron.2012.10.014).
- [53] L. Su, E. Fonteneau, W. Marslen-Wilson, N. Kriegeskorte. Spatiotemporal searchlight representational similarity analysis in EMEG source space. In *Proceedings of the 2nd International Workshop on Pattern Recognition in NeuroImaging*, IEEE, London, UK, pp. 97–100, 2012. DOI: [10.1109/PRNI.2012.26](https://doi.org/10.1109/PRNI.2012.26).
- [54] E. A. Wasserman, A. Chakroff, R. Saxe, L. Young. Illuminating the conceptual structure of the space of moral violations with searchlight representational similarity analysis. *NeuroImage*, vol. 159, pp. 371–387, 2017. DOI: [10.1016/j.neuroimage.2017.07.043](https://doi.org/10.1016/j.neuroimage.2017.07.043).
- [55] M. V. Peelen, A. Caramazza. Conceptual object representations in human anterior temporal cortex. *Journal of Neuroscience*, vol. 32, no. 45, pp. 15728–15736, 2012. DOI: [10.1523/JNEUROSCI.1953-12](https://doi.org/10.1523/JNEUROSCI.1953-12).
- [56] D. J. Kravitz, C. S. Peng, C. I. Baker. Real-world scene representations in high-level visual cortex: It's the spaces more than the places. *Journal of Neuroscience*, vol. 31, no. 20, pp. 7322–7333, 2011. DOI: [10.1523/JNEUROSCI.4588-10.2011](https://doi.org/10.1523/JNEUROSCI.4588-10.2011).
- [57] G. Handjaras, E. Ricciardi, A. Leo, A. Lenci, L. Cecchetti, M. Cosottini, G. Marotta, P. Pietrini. How concepts are encoded in the human brain: A modality independent, category-based cortical organization of semantic knowledge. *NeuroImage*, vol. 135, pp. 232–242, 2016. DOI: [10.1016/j.neuroimage.2016.04.063](https://doi.org/10.1016/j.neuroimage.2016.04.063).
- [58] D. I. Tamir, M. A. Thornton, J. M. Contreras, J. P. Mitchell. Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 1, pp. 194–199, 2016. DOI: [10.1073/pnas.1511905112](https://doi.org/10.1073/pnas.1511905112).
- [59] R. S. Chavez, T. F. Heatherton. Representational similarity of social and valence information in the medial pFC. *Journal of Cognitive Neuroscience*, vol. 27, no. 1, pp. 73–82, 2015. DOI: [10.1162/jocn_a_00697](https://doi.org/10.1162/jocn_a_00697).
- [60] B. Thirion, E. Duchesnay, E. Hubbard, J. Dubois, J. B. Poline, D. Lebihan, S. Dehaene. Inverse retinotopy: Inferring the visual content of images from brain activation patterns. *NeuroImage*, vol. 33, no. 4, pp. 1104–1116, 2006. DOI: [10.1016/j.neuroimage.2006.06.062](https://doi.org/10.1016/j.neuroimage.2006.06.062).
- [61] Y. Miyawaki, H. Uchida, O. Yamashita, M. A. Sato, Y. Morito, H. C. Tanabe, N. Sadato, Y. Kamitani. Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. *Neuron*, vol. 60, no. 5, pp. 915–929, 2008. DOI: [10.1016/j.neuron.2008.11.004](https://doi.org/10.1016/j.neuron.2008.11.004).
- [62] T. Naselaris, R. J. Prenger, K. N. Kay, M. Oliver, J. L.

- Gallant. Bayesian reconstruction of natural images from human brain activity. *Neuron*, vol. 63, no. 6, pp. 902–915, 2009. DOI: [10.1016/j.neuron.2009.09.006](https://doi.org/10.1016/j.neuron.2009.09.006).
- [63] S. Nishimoto, A. T. Vu, T. Naselaris, Y. Benjamini, B. Yu, J. L. Gallant. Reconstructing visual experiences from brain activity evoked by natural movies. *Current Biology*, vol. 21, no. 19, pp. 1641–1646, 2011. DOI: [10.1016/j.cub.2011.08.031](https://doi.org/10.1016/j.cub.2011.08.031).
- [64] Y. Fujiwara, Y. Miyawaki, Y. Kamitani. Modular encoding and decoding models derived from Bayesian canonical correlation analysis. *Neural Computation*, vol. 25, no. 4, pp. 979–1005, 2013. DOI: [10.1162/NECO_a_00423](https://doi.org/10.1162/NECO_a_00423).
- [65] A. S. Cowen, M. M. Chun, B. A. Kuhl. Neural portraits of perception: Reconstructing face images from evoked brain activity. *NeuroImage*, vol. 94, pp. 12–22, 2014. DOI: [10.1016/j.neuroimage.2014.03.018](https://doi.org/10.1016/j.neuroimage.2014.03.018).
- [66] C. D. Du, C. Y. Du, H. G. He. Sharing deep generative representation for perceived image reconstruction from human brain activity. In *Proceedings of International Joint Conference on Neural Networks*, IEEE, Anchorage, USA, pp. 1049–1056, 2017. DOI: [10.1109/IJCNN.2017.7965968](https://doi.org/10.1109/IJCNN.2017.7965968).
- [67] D. P. Kingma, M. Welling. Auto-encoding variational Bayes. [Online], Available: <https://arxiv.org/abs/1312.6114>, 2014.
- [68] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, Y. Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal, Canada, pp. 2672–2680, 2014.
- [69] T. Horikawa, Y. Kamitani. Generic decoding of seen and imagined objects using hierarchical visual features. *Nature Communications*, vol. 8, Article number 15037, 2017. DOI: [10.1038/ncomms15037](https://doi.org/10.1038/ncomms15037).
- [70] G. St-Yves, T. Naselaris. Generative adversarial networks conditioned on brain activity reconstruct seen images. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, IEEE, Miyazaki, Japan, pp. 1054–1061, 2018. DOI: [10.1109/SMC.2018.00187](https://doi.org/10.1109/SMC.2018.00187).
- [71] Y. Güçlütürk, U. Güçlü, K. Seeliger, S. Bosch, R. van Lier, M. van Gerven. Reconstructing perceived faces from brain activations with deep adversarial neural decoding. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, USA, pp. 4249–4260, 2017.
- [72] K. Seeliger, U. Güçlü, L. Ambrogioni, Y. Güçlütürk, M. A. J. van Gerven. Generative adversarial networks for reconstructing natural images from brain activity. *NeuroImage*, vol. 181, pp. 775–785, 2018. DOI: [10.1016/j.neuroimage.2018.07.043](https://doi.org/10.1016/j.neuroimage.2018.07.043).
- [73] R. VanRullen, L. Reddy. Reconstructing faces from fMRI patterns using deep generative neural networks. *Communications Biology*, vol. 2, no. 1, Article number 193, 2019. DOI: [10.1038/s42003-019-0438-y](https://doi.org/10.1038/s42003-019-0438-y).
- [74] R. Belyi, G. Gaziv, A. Hoogi, F. Strappini, T. Golan, M. Irani. From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vancouver, Canada, 2019.
- [75] D. Li, C. D. Du, H. G. He. Semi-supervised cross-modal image generation with generative adversarial networks. *Pattern Recognition*, vol. 100, Article number 107085, 2020. DOI: [10.1016/j.patcog.2019.107085](https://doi.org/10.1016/j.patcog.2019.107085).
- [76] C. D. Du, C. Y. Du, H. Wang, J. P. Li, W. L. Zheng, B. L. Lu, H. G. He. Semi-supervised deep generative modelling of incomplete multi-modality emotional data. In *Proceedings of the 26th ACM international conference on Multimedia*, ACM, Seoul, Republic of Korea, pp. 108–116, 2018. DOI: [10.1145/3240508.3240528](https://doi.org/10.1145/3240508.3240528).
- [77] X. Cai, F. P. Nie, W. D. Cai, H. Huang. Heterogeneous image features integration via multi-modal semi-supervised learning model. In *Proceedings of IEEE International Conference on Computer Vision*, IEEE, Sydney, Australia, 2013. DOI: [10.1109/ICCV.2013.218](https://doi.org/10.1109/ICCV.2013.218).
- [78] Z. X. Zhang, F. Ringeval, B. Dong, E. Coutinho, B. Schuller. Enhanced semi-supervised learning for multimodal emotion recognition. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, Shanghai, China, 2016. DOI: [10.1109/ICASSP.2016.7472666](https://doi.org/10.1109/ICASSP.2016.7472666).
- [79] T. Schonberg, C. R. Fox, J. A. Mumford, E. Congdon, C. Trepel, R. A. Poldrack. Decreasing ventromedial prefrontal cortex activity during sequential risk-taking: An fMRI investigation of the balloon analog risk task. *Frontiers in Neuroscience*, vol. 6, Article number 80, 2012. DOI: [10.3389/fnins.2012.00080](https://doi.org/10.3389/fnins.2012.00080).
- [80] A. R. Aron, M. A. Gluck, R. A. Poldrack. Long-term test-retest reliability of functional MRI in a classification learning task. *NeuroImage*, vol. 29, no. 3, pp. 1000–1006, 2006. DOI: [10.1016/j.neuroimage.2005.08.010](https://doi.org/10.1016/j.neuroimage.2005.08.010).
- [81] S. M. Tom, C. R. Fox, C. Trepel, R. A. Poldrack. The neural basis of loss aversion in decision-making under risk. *Science*, vol. 315, no. 5811, pp. 515–518, 2007. DOI: [10.1126/science.1134239](https://doi.org/10.1126/science.1134239).
- [82] K. Foerde, B. J. Knowlton, and R. A. Poldrack. Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 31, pp. 11778–11783, 2006. DOI: [10.1073/pnas.0602659103](https://doi.org/10.1073/pnas.0602659103).
- [83] R. A. Poldrack, J. Clark, E. J. Paré-Blagoev, D. Shohamy, J. C. Moyano, C. Myers, M. A. Gluck. Interactive memory systems in the human brain. *Nature*, vol. 414, no. 6863, pp. 546–550, 2001. DOI: [10.1038/35107080](https://doi.org/10.1038/35107080).
- [84] A. M. C. Kelly, L. Q. Uddin, B. B. Biswal, F. X. Castellanos, M. P. Milham. Competition between functional brain networks mediates behavioral variability. *NeuroImage*, vol. 39, no. 1, pp. 527–537, 2008. DOI: [10.1016/j.neuroimage.2007.08.008](https://doi.org/10.1016/j.neuroimage.2007.08.008).
- [85] K. J. Duncan, C. Pattamadilok, I. Knierim, J. T. Devlin. Consistency and variability in functional localisers. *NeuroImage*, vol. 46, no. 4, pp. 1018–1026, 2009. DOI: [10.1016/j.neuroimage.2009.03.014](https://doi.org/10.1016/j.neuroimage.2009.03.014).
- [86] J. M. Walz, R. I. Goldman, M. Carapezza, J. S. Muraskin, T. R. Brown, P. Sajda. Simultaneous EEG-fMRI reveals temporal evolution of coupling between supramodal cortical attention networks and the brainstem. *Journal of Neuroscience*, vol. 33, no. 49, pp. 19212–19222, 2013. DOI: [10.1523/JNEUROSCI.2649-13.2013](https://doi.org/10.1523/JNEUROSCI.2649-13.2013).
- [87] T. D. Verstynen. The organization and dynamics of corticostriatal pathways link the medial orbitofrontal cortex to future behavioral responses. *Journal of Neurophysiology*, vol. 112, no. 10, pp. 2457–2469, 2014. DOI: [10.1152/jn.00221.2014](https://doi.org/10.1152/jn.00221.2014).
- [88] M. G. Veldhuizen, R. K. Babbs, B. Patel, W. Fobbs, N. B. Kroemer, E. Garcia, M. R. Yeomans, D. M. Small. Integration of sweet taste and metabolism determines carbohydrate reward. *Current Biology*, vol. 27, no. 16, pp. 2476–2485.e6, 2017. DOI: [10.1016/j.cub.2017.07.018](https://doi.org/10.1016/j.cub.2017.07.018).
- [89] M. Hanke, Y. O. Halchenko, P. B. Sederberg, S. J. Hanson, J. V. Haxby, S. Pollmann. PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroin-*

formatics, vol.7, no.1, pp.37–53, 2009. DOI: [10.1007/s12021-008-9041-y](https://doi.org/10.1007/s12021-008-9041-y).

- [90] M. Hanke, Y. O. Halchenko, P. B. Sederberg, E. Olivetti, I. Fründ, J. W. Rieger, C. S. Herrmann, J. V. Haxby, S. JoséHanson, S. Pollmann. PyMVPA: A unifying approach to the analysis of neuroscientific data. *Frontiers in Neuroinformatics*, vol.3, Article number 3, 2009. DOI: [10.3389/neuro.11.003.2009](https://doi.org/10.3389/neuro.11.003.2009).
- [91] H. J. Zhang, P. H. Chen, P. Ramadge. Transfer learning on fMRI datasets. In *Proceedings of the 21st International Conference on Artificial Intelligence and Statistics*, Playa Blanca, Spain, pp. 595–603, 2018.
- [92] M. L. Wang, D. Q. Zhang, J. S. Huang, P. T. Yap, D. G. Shen, M. X. Liu. Identifying autism spectrum disorder with multi-site fMRI via low-rank domain adaptation. *IEEE Transactions on Medical Imaging*, vol.39, no.3, pp.644–655, 2020. DOI: [10.1109/TMI.2019.2933160](https://doi.org/10.1109/TMI.2019.2933160).
- [93] X. Zhang, Q. Yang. Transfer hierarchical attention network for generative dialog system. *International Journal of Automation and Computing*, vol.16, no.6, pp.720–736, 2019. DOI: [10.1007/s11633-019-1200-0](https://doi.org/10.1007/s11633-019-1200-0).



Shuo Huang received the B.Sc. degree in software engineering from Northeastern University, China in 2015. He is currently a Ph.D. degree candidate in software engineering in College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics (NUAA), China.

His research interests include machine learning and human brain decoding.

E-mail: huangshuo@nuaa.edu.cn

ORCID iD: 0000-0002-3267-1816



Wei Shao received the B.Sc. and M.Sc. degrees in information and computing science from Nanjing University of Technology, China in 2009 and 2012, respectively, and the Ph.D. degree in software engineering from Nanjing University of Aeronautics and Astronautics, China in 2018.

His research interests include machine learning and bioinformatics.

E-mail: 527606857@qq.com

ORCID iD: 0000-0003-1476-2068



Mei-Ling Wang received the M.Sc. degree in information and communication engineering from Nanjing University of Information Science and Technology, China in 2016. She is currently a Ph.D. degree candidate in College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China.

Her research interests include machine learning and brain imaging genetics.

E-mail: mely@nuaa.edu.cn

ORCID iD: 0000-0001-6569-2798



Dao-Qiang Zhang received the B.Sc. and Ph.D. degrees in computer science from Nanjing University of Aeronautics and Astronautics, China in 1999, and 2004, respectively. He joined Department of Computer Science and Engineering of NUAA as a lecturer in 2004, and is a professor at present. He has published over 200 scientific articles in refereed international

journals such as *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *IEEE Transactions on Medical Imaging*, *IEEE Transactions on Image Processing*, *Neuroimage*, *Human Brain Mapping*, *Medical Image Analysis*; and conference proceedings such as IJCAI, AAAI, NIPS, CVPR, MICCAI, KDD, with 12000+ citations by Google Scholar. He was nominated for the National Excellent Doctoral Dissertation Award of China in 2006, won the Best Paper Award and the Best Student Award of several international conferences such as PRICAI'06, STMI'12 and BICS'16, etc. He has served as a program committee member for several international conferences such as IJCAI, AAAI, NIPS, MICCAI, SDM, PRICAI, ACML, etc. He is a member of the Machine Learning Society of the Chinese Association of Artificial Intelligence (CAAI), and the Artificial Intelligence & Pattern Recognition Society of the China Computer Federation (CCF).

His research interests include machine learning, pattern recognition, data mining and medical image analysis.

E-mail: dqzhang@nuaa.edu.cn (Corresponding author)

ORCID iD: 0000-0002-5658-7643

Citation: S. Huang, W. Shao, M. L. Wang, D. Q. Zhang. Fmri-based decoding of visual information from human brain activity: a brief review. *International Journal of Automation and Computing*, vol.18, no.2, pp.170–184, 2021. <https://doi.org/10.1007/s11633-020-1263-y>

Articles may interest you

An advanced analysis system for identifying alcoholic brain state through eeg signals. *International Journal of Automation and Computing*, vol.16, no.6, pp.737-747, 2019.

DOI: [10.1007/s11633-019-1178-7](https://doi.org/10.1007/s11633-019-1178-7)

A survey on 3d visual tracking of multicopters. *International Journal of Automation and Computing*, vol.16, no.6, pp.707-719, 2019.

DOI: [10.1007/s11633-019-1199-2](https://doi.org/10.1007/s11633-019-1199-2)

Selection of observation position and orientation in visual servoing with eye-in-vehicle configuration for manipulator. *International Journal of Automation and Computing*, vol.16, no.6, pp.761-774, 2019.

DOI: [10.1007/s11633-019-1181-z](https://doi.org/10.1007/s11633-019-1181-z)

Electronic nose and its applications: a survey. *International Journal of Automation and Computing*, vol.17, no.2, pp.179-209, 2020.

DOI: [10.1007/s11633-019-1212-9](https://doi.org/10.1007/s11633-019-1212-9)

Deep learning based single image super-resolution: a survey. *International Journal of Automation and Computing*, vol.16, no.4, pp.413-426, 2019.

DOI: [10.1007/s11633-019-1183-x](https://doi.org/10.1007/s11633-019-1183-x)

Transfer hierarchical attention network for generative dialog system. *International Journal of Automation and Computing*, vol.16, no.6, pp.720-736, 2019.

DOI: [10.1007/s11633-019-1200-0](https://doi.org/10.1007/s11633-019-1200-0)

Expression analysis based on face regions in real-world conditions. *International Journal of Automation and Computing*, vol.17, no.1, pp.96-107, 2020.

DOI: [10.1007/s11633-019-1176-9](https://doi.org/10.1007/s11633-019-1176-9)

The propagation background in social networks: simulating and modeling. *International Journal of Automation and Computing*, vol.17, no.3, pp.353-363, 2020.

DOI: [10.1007/s11633-020-1227-2](https://doi.org/10.1007/s11633-020-1227-2)



WeChat: IJAC



Twitter: IJAC_Journal